METHOD FOR THE DETECTION OF ABNORMALLY GLYCOSYLATED PROTEINS

The present invention relates to a method of screening a sample for the presence of one or more abnormally glycosylated and/or expressed proteins. The present invention is particularly, although not exclusively, concerned with screening a biological sample, for example a body fluid sample, for the presence of abnormally O- and/or N-glycosylated proteins, which may be indicative of disease or of substance abuse in an individual.

Most plasma membrane and secretory proteins are complex macromolecules incorporating oligosaccharides, which play an integral role in their biological function. Consequently, abnormal glycosylation or expression of proteins is often associated with disease. For example, defects in nine genes of the N-linked glycosylation pathway are associated with serious medical conditions, including neurological dysfunction, collectively termed Congenital Disorders of Glycosylation (CDG, H. Freeze, Glycobiology 2001, 11, 129R-143R). Other conditions associated with abnormal glycosylation include Leroy disease, Wiskott-Aldrich syndrome and glycoproteinosis. Further, diseases resulting from exposure to some external factor, such as alcoholism, rheumatoid arthritis and cancer are also associated with abnormal glycosylation of proteins (see for example, G. Durand and N. Seta, Clin. Chem. 2000, 46, 795-805).

In addition, it is known (V. Skibeli et al., Blood 2001, 98, 3626-3634) that the glycosylated form of recombinant erythropoietin (rEPO), which is used by athletes to boost performance, differs from the natural form in that it is deficient in one type of sialic acid linkage (Neu5Ac(α2-6)Gal on N-glycans) and has supplementary poly-N-

acetyl lactosamine (Gal(β1-4)GlcNAc(b1-3)) and glucosamine (GlcNAc(β1-4)Man(β1-4)). Exogenous erythropoietin can therefore be distinguished from the endogenous protein.

Presently, methods for determining abnormal glycosylation rely on iso-electric focusing (IEF) or Western blotting of blood extracts for different glycoforms of human serum transferin (H. Stibler et al., Acta. Paediatr. Scand. Suppl., 1991, 375, 21-31; J. Jaeken et al., Clin, Chim. Acta, 1984, 144, 245 to 247; N. Seta et al., Clin. Chim. Acta 1996, 24, 131-140). Recently, however, abnormal glycosylation has been determined by lectin analysis of purified serum endocrine glycoproteins (M. Ferrari et al., Eur. J. Endocrinol., 2001, 144, 409-416).

The detection of illicit use of rEPO is currently based on IEF of urine concentrates followed by a double blotting procedure using antibodies (F. Lasne and J. de Ceaurriz, Nature 2000, 405, 635 and F. Lasne, J. Immunol. Methods, 2001, 253 (1-2), 125-131).

However, all of the above techniques require preliminary concentration, purification or isolation of a predetermined protein and can be time consuming and/or expensive. IEF involves a multitude of steps and can often take up to three days. Further, it is not specific in that it is based on net charge of the molecule so that different defects, leading to identical net charge can have substantially similar focusing patterns.

International Patent Application WO 00/33076 discloses a method of diagnosis of human glycosylation disorders. A first, reagent, for example a lectin, binds with an oligosaccharide moiety present in a sample from an individual having a glycosylation

disorder and not present in a sample from a healthy individual. A second, reagent binds with an oligosaccharide present in a healthy individual but not present in an individual having a glycosylation disorder. A glycosylation disorder is indicated through binding of the first reagent and/or non-binding of the second reagent.

The method is, however, dependent on the presence or absence of a target oligosaccharide moiety in the sample compared with a control sample. It is, therefore, unable to detect glycosylation disorders characterised by the relative extent and/or ratio of oligosaccharide moieties (or epitopes). There is consequently a need for an improved method for determining abnormal glycosylation or substance abuse in individuals.

The present invention generally seeks to provide an improved method for screening a sample, in particular a body fluid sample, for one or more abnormally glycosylated proteins. The present invention also aims to improve diagnosis of disease exhibiting abnormal glycosylation and to improve testing for performance enhancing protein drugs in mammals.

The present invention also seeks to provide an improved method for determining the presence of an abnormally expressed protein in a body fluid sample. Such a method may be of value to the diagnosis of disease, such as Alzeihemer's, BSE and the like.

Accordingly, in a first aspect, the present invention provides a method of screening a sample for the presence of one or more abnormally glycosylated and/or expressed proteins comprising the steps of i) exposing said sample to two or more different

lectins and/or antibodies ii) determining the extent of binding of said sample to at least two of said lectins and/or antibodies and iii) comparing the determined extent of binding to said at least two lectins and/or antibodies with that of a control sample.

The method is preferably used for screening biological samples, i.e. those taken from the human or animal body. However, it may also be used to screen other samples such as those used during laboratory experimentation.

It will be understood that a particular sample may be exposed to the lectins and/or antibodies separately and sequentially. In a preferred embodiment, however, the sample is exposed to the said two or more lectins and/or antibodies simultaneously.

In particular, the lectins and/or antibodies may be immobilised on a solid support surface and arranged in an array and the sample sequentially and/or simultaneously exposed to them.

Although the determination of the extent of binding to a selection of different lectins and/or antibodies is of critical importance to the present invention, there is no requirement that more than two, or any one particular lectin or antibody is used. All that is required is that a pattern of the extent of binding is determined for a particular sample. Preferably, however, the extent of binding is determined for more than two lectins and/or antibodies.

Suitable lectins include those that are specific for sialic acid (Neu5Ac), galactose, mannose, glucosamine and fucose containing oligosaccharides. Amongst these, the

commercially available lectins *Sambucus nigra* agglutinin (SNA), specific for α-2, 6–linked Neu5Ac and *Maackia amurensis* agglutinin (MAA), specific for α-2, 3–linked Neu5Ac are preferred. Other suitable, lectins include *Tetragonolobus purpureaus* (TLP), *Anguila anguila* (AA) and *Ulex europaeus I* agglutinin (UEA) which are specific for Fuc(α1-3/4). Suitable antibodies may be raised by any technique known to the art.

As mentioned above, the extent of binding of a particular sample to a lectin and/or antibody array, for example, follows a unique pattern. The pattern, which can be thought of as a "fingerprint" or "signature" of the sample on the array, characterises the glycosylation environment of the sample, each lectin, for example, being specific to a group of related oligosaccharide epitopes.

The fingerprint or signature of a particular sample may be expressed by one or more ratios of extent of binding to different lectins and/or antibodies. Thus, in a preferred embodiment the comparison with the control sample compares one or more ratios (A/B, A/C, B/C etc.) of extent of binding to different lectins and/or antibodies (A, B, C) for each sample.

It will be apparent, therefore, that the method is particularly suited to screening complex samples - which may be expected to exhibit binding to most or all of the lectins and/or antibodies.

The method can avoid the need for concentration, purification and isolation of a previously identified protein since, normally, it is the comparison of the fingerprint or

signature with that of a control sample that is indicative of the presence of an abnormally glycosylated or expressed protein. Thus, in one embodiment of the present invention, a body fluid sample may be directly exposed to the array.

Of course, the selection of a particular body fluid for screening will be based on the amount of glycoprotein supposed or known to be directly available therein and the sensitivity of the method utilised in the determination step. Suitable body fluids for direct analysis include urine and blood serum.

It may, however, be desirable to concentrate or purify the sample. In another embodiment, therefore, the method of the present invention comprises the preliminary step of isolating one or more proteins from the sample by capture with an antibody or antibodies. In this embodiment, the isolation of protein or proteins minimises unwanted complications in the determination step – especially where the concentration of the proteins approaches the limits of detection and quantification.

The fingerprint or signature of a particular sample may be interrogated by any suitable means determining the extent of binding of the sample to said at least two lectins and/or antibodies. In particular, the extent of binding of a particular sample to a particular lectin or antibody may be determined by colorimetric, fluorescence or optical techniques.

In a preferred embodiment, however, that the fingerprint or signature is interrogated by a technique allowing real-time monitoring of the binding of the sample to the array. Such techniques also allow kinetic characteristics such as binding or

dissociation rates, to be determined providing further information as to the presence of abnormal glycosylated or expressed proteins.

In particular, the fingerprint or signature may be interrogated by an evanescent optical technique. Preferably, therefore, the solid support and array together comprise an optical biosensor. Suitable solid supports include surface plasmon resonance (SPR) chips and/or metal or dye clad leaky waveguide (MCLW/DCLW) chips, such as those described in our international applications WO 99/44042 and PCT/GB02/045045.

Preferably, the optical interrogation utilises a flow technique. Suitable apparatus, therefore, include ordinary and optimised surface plasmon resonance apparatus described in International Patent Application WO 01/42768.

The comparison of the interrogated fingerprint or signature with that of the control sample does not necessarily require that the investigator himself make an exposure of the control sample to an identical array.

In particular, the comparison may be made by reference to a library database including fingerprint data from a number of control samples. Preferably, the comparison is computer aided. In this embodiment, the fingerprint data deposited in the database need only record salient features characterising health and/or abnormal glycosylation.

An exposure of the control sample may, however, be preferred in lectin-only arrays where the relatively weak lectin-glycoprotein interaction allows easy regeneration.

The control sample may be derived from a healthy individual or an individual exhibiting abnormal glycosylation. In particular, the control sample may be one or more earlier samples taken from the same individual. Such a control sample is especially useful for monitoring continuation of substance abuse or the progress of disease and/or treatment.

It will appreciated from the foregoing, that the present invention provides an improved method of screening samples, in particular body fluid samples, in that it compares the ratio of binding to two or more lectins and/or antibodies. Thus the method is not dependent on the absence of certain epitopes in one or other of the samples and is less prone to false positives.

In addition, real-time monitoring based on optical techniques allows rapid screening of complex samples in a single exposure of the sample to a lectin and/or antibody array. Such techniques allow simultaneous determination of binding to each lectin and/or antibody in an array and may offer additional information which may characterise the absence or presence of an abnormally glycosylated or expressed protein in a sample.

The method of the present invention may also be used to screen a body fluid sample for indication of use or abuse of glycoprotein drugs in humans or animals.

In a second aspect, therefore, the present invention provides a method for determining use of a glycoprotein drug in a human or animal comprising the steps of i) taking a body fluid sample ii) exposing said sample to two or more different lectins and/or

antibodies iii) determining the extent of binding of said sample to at least two lectins and/or antibodies and iii) comparing the determined extent of binding to said at least two lectins and/or antibodies with that of a control sample.

It will be understood that the glycoprotein drug is synthetic in origin and differs from the endogenous form in, for example, its glycosylation environment. The method is particularly suitable for determining abuse or illicit use of recombinant performance enhancing protein drugs, such as recombinant forms of erythropoietin, chorionic gonadotropin or human growth hormone.

The method provides for detection of exogenous protein through pattern recognition indicating the presence of abnormal glycoforms of the protein - even in the presence of endogenous protein.

Exposure of a urine sample to a lectin-only array, for example, and comparison of the interrogated fingerprint in relation and/or other indicia with that of a control sample as explained above can provide conclusive evidence of use.

Embodiments of the second aspect of the present invention will be apparent from the above and the claims appended hereto.

As mentioned above, the fingerprint or signature of a particular sample may be informative as to a particular disease state in an individual when compared to that of a sample from a healthy individual.

In a third aspect, therefore, the present invention provides a method for the diagnosis of acquired or inherited glycosylation disorders comprising the steps of i) taking a body fluid sample ii) exposing said sample to two or more different lectins and/or antibodies iii) determining the extent of binding of said sample to at least two lectins and/or antibodies and iii) comparing the determined extent of binding to said at least two lectins and/or antibodies with that of a control sample.

Embodiments of the third aspect of the present invention will also be apparent from the above and the claims appended hereto.

The method in this aspect of the present invention may, for example, be particularly suitable for diagnosis of disease in neonates, children or adults. The method is not necessarily limited to medical conditions which are attributable to abnormal glyscosylation but may also be of value to conditions accompanied by abnormal glycosylation.

In a further aspect, the present invention provides a kit of parts for use in the aforementioned methods comprising one or more lectins and/or antibodies and a control sample and/or information relating to normal or expected glycosylation binding patterns and/or characteristics for an individual type.

The lectins and/or antibodies may be supplied in vials or a suitable container or containers. Additionally, a solid support surface may be supplied for attachment of a selection of lectins and/or antibodies by any of the usual methods known to the art. Preferably, the solid support surface comprises an SPR, MCLW or DCLW chip.

In a preferred embodiment, suitable for detecting the illicit use of a performance enhancing drug, such as erythropoietin, the kit may comprise only lectins.

The supplied information may also include information relating binding patterns and/or characteristics to candidate disease states. In particular, the additional information may categorise certain binding patterns and/or characteristics according to likely medical condition.

The present invention will now be described by reference to the following examples and drawings in which

Figure 1 is a putative bar graph illustrating the fingerprint or signature of samples (a, b, c) on a four-fold lectin array;

Figures 2 a) and b) are SPR sensor-grams illustrating respectively the binding of buffered rEPO and bovine fetuin solutions to a SNA/MAA array;

Figures 3 a), b) and c) are SPR sensor-grams illustrating respectively the binding of a buffered solution of NESP (a rEPO), human EPO and human fetuin to a SNA/MAA array;

Figure 4 is a bar graph showing a fingerprint or signature for the samples of Figures 3 a) to c);

Figures 5 a) to d) are SPR sensor-grams illustrating the binding of buffered solutions of mixtures of human EPO/NESP and human EPO/human fetuin on a SNA/MAA array;

Figure 6 is a graph showing the dissociation characteristics of human EPO/NESP mixtures from MAA according to Figure 5 relative to the proportion of EPO;

Figures 7 a) and b) are graphs showing the ratio of response to the SNA/MAA array according to Figure 5 of human EPO/NESP mixtures and human EPO/human fetuin mixtures relative to the proportion of EPO;

Figure 8 is a SPR sensor-gram illustrating the interaction of a urine sample respectively with the SNA/MAA array; and

Figures 9 a) and b) are graphs highlighting the screening of human blood serum samples for alcoholism or related diseases using a four-fold lectin array.

Referring now to Figure 1, the putative response of four lectins (1 to 4) to a control sample a) and two samples b) and c) exhibiting an abnormal glycosylation disorder confer a fingerprint or signature to each sample (a1, a2, a3, a4 etc). The fingerprint or signature of sample b), when compared to that of the control sample, indicates the disorder although it is not apparent from its interaction with lectin 2 alone.

Having regard to Figures 2 to 8, experiments were performed on a BIAcore® 2000 four flow channel SPR instrument equipped with BIAevaluation software 3.0. Lectins *Sambucus nigra* agglutinin and *Maackia amurensis* agglutinin (Sigma, Madrid) were each immobilised to part of an SPR chip (carboxymethylated dextran surface, standard density CM5 sensorchip, Pharmacia Biosensor AB, Uppsala, Sweden) for registration with each channel using known amine coupling techniques (S.R. Haseley et al., Anal. Biochem., 1999, 271, 203-210). Table 1 summarises the immobilisation parameters for the chip.

Channel (Ch)	Ch 1	Ch 2	Ch 3
Immobilisation Buffer (IB)	None	NaOAc, 10 mM, pH 4.5	NaOAc, 10 mM, pH 4.5
Lectin Solution	Blank (chemically activated for lectin immobilisation)	4 µl of 5 $mgml^{-1}$ MAA solution in H_2O diluted to 100 µl IB	10 µl of 2 $mgml^{-1}$ SNA solution in H_20 diluted to 100 µl IB
Immobilisation (Response Units RU)	40 RU	9000 RU	14000RU

Table 1

Example 1

rEPO expressed in CHO cells (European Pharmacopeia, France) was made up to 5 μgml^{-1} in distilled water. Bovine fetuin (Sigma, Madrid) was made up to $10\mu gml^{-1}$ in distilled water. The interactions of 35 µl of each of the rEPO and bovine fetuin solution with the chip was studied in running buffer (pH 7.5; 100 mM HEPES, 150 mM NaCl, 2 mM $CaCl_2$ and 2 mM $MgCl_2$ in distilled water).

Referring now to Figure 2 a), the binding of rEPO to MAA is much greater than to SNA. rEPO, in common with the natural form of EPO, has a significant number of α-2,3-Neu5Ac linkages - specifically recognised by MAA but, in contrast to the natural form of EPO, only a single α-2,6-Neu5Ac linkage - specifically recognised by SNA.

The binding of rEPO with MAA shows a rapid uptake of the glycoprotein to the lectin, which gradually levels off. A low level of disassociation of the glycoprotein follows - the residual level of binding remaining high.

The binding of rEPO with SNA include an initial uptake of the glycoprotein, which almost immediately levels off. Residual binding of the glycoprotein tends to zero even though it contains one α-2,6-Neu5Ac linkage.

Referring now to Figure 2 b), the binding of bovine fetuin, which is similar to human EPO in that it comprises significant numbers of both α-2,6-Neu5Ac and α-2,3-Neu5Ac linkages, to MAA is gradual and low. By contrast, the binding to SNA is rapid and sustained at high levels.

It will be appreciated that concentration effects and the presence of linkages in common do not permit one or other of these lectins alone to definitively indicate whether the sample contains exogenous protein in addition to the endogenous protein. However, these results suggest that the interaction of a sample with two or more lectins will, however, allow exogenous protein to be distinguished - even in the presence of the natural form.

Further, the presence of rEPO in a sample will be apparent not just through analysis of the ratio of MAA/SNA binding compared with that from a healthy mammal but also through comparison of the analysis of uptake, equilibrium and/or dissociation characteristics of each sample.

Example 2

Referring now to Figures 3 a) to c), the binding (after 7 min. of 5 μl min^{-1} running buffer TRIS, pH 7.2, 150 mM NaCl, 2 mM $CaCl_2$ and 2mM $MgCl_2$) of each of human EPO (1 μl from stock in 200 μl, b) and human fetuin (7.5 $\mu g ml^{-1}$ solution, c), to a

similar chip is greater to SNA than to MAA. However, the binding of NESP (8 µl from stock in 200 µl, a), a recombinant form of EPO, to SNA is almost negligible when compared to MAA.

Figure 4 shows the fingerprint or signature of each sample in relation to both lectins. As may be seen the ratio of binding having regard to both lectins distinguishes the samples from each other.

Referring now to Figure 5, the interaction of each lectin with mixtures a) to c) of varying ratios of human EPO to NESP, are of interest not only in the ratio of the level of binding but also in the dissociation phase of each response. It will be noted that the fingerprints and the dissociation curves are different for each mixture as well as for the pure samples.

The linear relationship between the amount of EPO and k_d for MAA and/or SNA as well as SNA/MAA binding ratio also suggests that the method may, in combination with other techniques, enable quantitative determination of rEPO in a sample.

Table 2 compares the ratio of response of the lectins and the rate of dissociation k_d for each sample. As is clear, especially in regard to Figures 6 and 7 a), the ratio of binding response SNA/MAA and the rate of dissociation from MAA varies linearly with the amount of EPO in the mixture. These results suggest that NESP and, indeed, human fetuin can be determined in the presence of human EPO.

Sample	Binding Response (10 s > dissociation point)		Ratio of Residual Binding Responses		$k_d/10^{-4}$	
	SNA	MAA	SNA/MAA	MAA/SNA	SNA	MAA
EPO (E)	234	160	1.463	0.684	3.30	3.09
NESP (N)	0	157	0	-	-	-3.27
1:1 E/N	215	282	0.762	1.312	2.16	0.05
3:1 E/N	401	361	1.111	0.900	2.81	2.08
1:2 E/N	200	305	0.665	1.525	1.34	0.29
Fetuin (F)	245	18	13.611	0.073	0.70	-7.29
1:1 E/F	307	138	2.225	0.450	1.07	1.88

Table 2

Sensitivity experiments performed on the chip showed that the lower detection limit of rEPO by MAA was in the range of 5 ngml-1 (160 pmol.l^{-1}) which is comparable to the reference solution used for IEF (380 pmol.l^{-1}).

Referring now to Figure 8, the results of the interaction of a control urine sample, concentrated 20-fold, is shown. Here, the chip has in addition to a blank (-x- line) surface, has a surface in which the dextran surface is not activated for lectin immobilisation (-+-). As may be seen, glycoprotein interactions with MAA and SNA are present.

Reference to values (0 to 5.3 $pmol.l^{-1}$) published for EPO in urine suggest the need for a 200 fold increase in sensitivity. It is envisaged, however, that routine optimisation of the technique (by selection of lectins and running buffers) and/or detection method (for example, light or emitted from the array through fluorescent tagging) will enable urine samples to be directly screened without the need for a concentration or isolation step.

For the determination of abuse of rEPO large arrays including MAA, SNA and lectins specific for poly N-acetyl lactosamine repeat units (Gal(β1-4)GlcNAc(β1-3) or lectins specific for bisecting GlcNAc (GlcNAc(β1-4)Man(β1-4) are expected to offer significant improvements. Alternatively or additionally, arrays including lectins or antibodies specific for tetraantennary glycan moieties found mostly with the recombinant form may also be of value.

Channel (Ch)	Ch 1	Ch 2	Ch 3	Ch 4
Lectin	BPA	MAA	RCA_{120}	SNA
Immobilisation (Response Units RU)*	9600	6500	1100	1100

Table 3

Having regard now to Figures 9 a) and b), similar experiments were performed on 24 human blood serum samples using an SPR chip including an array of four lectins *Bauhinia purpureus* agglutinin (BPA), *Maackia amurensis* agglutinin, *Sambucus*

nigra agglutinin and *Ricinus communis* 120 agglutinin (RCA_{120}). Table 3 summarises the immobilisation responses of each lectin on the chip.

Example 3

Half of the serum samples were obtained from patients with alcoholism or related diseases, and so contained high (H1-H12) levels of carbohydrate deficient transferrin (CDT) relative to the total content of transferrin. The other half of the serum (control) samples had normal (low, L1 to L12) levels of carbohydrate deficient transferrin (CDT) relative to the total content of transferrin.

The serum samples were simply diluted (1:100) in running buffer (10 mM Tris.HCl, pH 7.0; 150 mM NaCl; 2 mM $CaCl_2$ and 2 mM $MgCl_2$) and the interaction of each diluted sample with the chip was studied at various time (t) intervals

At t = 150 s the high (between 3.5 to 9.2 % transferrin) and low (between 2.0 to 2.6 % transferrin) CDT containing samples can be reliably distinguished by examining one or more of the order of response in the lectin array as well as the magnitude and relative magnitude of these responses.

For example, high CDT samples were found to have, in general, an order of response Ch: 3, 4, 1, 2 or 3, 2, 4, 1 whereas the majority of low CDT samples gave an order of response 3, 4, 2, 1.

The difference (D, 4-2) in response of SNA and MAA to high CDT samples was generally high or negative compared with low CDT samples. The ratio (R, 3/4) in

response of RCA120 to SNA to low CDT samples tended to be lower than high CDT samples.

A plot of difference D, 4-2 against R, 3/4 (Figure 9 a) shows tight clustering of the majority of low CDT samples and two distinct clusters of high CDT samples. A plot of ratio R, 3/4 against ratio R, 4/2 (Figure 9 b) similarly reveals tight clustering of low CDT samples.

The present invention may be suitable for rapid recognition of a wide variety of inherited or congenital disorders associated with abnormal glycosylation. Blood serum or urine may be screened for other medical conditions, not mentioned above.

In particular, it is envisaged that the method is suitable for prenatal screening for Down's syndrome in that concentrations of human chorionic gonadotropin (hCG) or sub-units thereof are high in the serum and urine of pregnant women.

It is also envisaged that a wide variety of other lectins may be used and that the interaction of a sample with large numbers of lectins will not be unduly complicated – especially where analysis is computer-aided.